

Composer: Creative and Controllable Image Synthesis with Composable Conditions

Year: 2023 | Citations: 347 | Authors: Lianghua Huang, Di Chen, Yu Liu, Yujun Shen, Deli Zhao

Abstract

Recent large-scale generative models learned on big data are capable of synthesizing incredible images yet suffer from limited controllability. This work offers a new generation paradigm that allows flexible control of the output image, such as spatial layout and palette, while maintaining the synthesis quality and model creativity. With compositionality as the core idea, we first decompose an image into representative factors, and then train a diffusion model with all these factors as the conditions to recompose the input. At the inference stage, the rich intermediate representations work as composable elements, leading to a huge design space (i.e., exponentially proportional to the number of decomposed factors) for customizable content creation. It is noteworthy that our approach, which we call Composer, supports various levels of conditions, such as text description as the global information, depth map and sketch as the local guidance, color histogram for low-level details, etc. Besides improving controllability, we confirm that Composer serves as a general framework and facilitates a wide range of classical generative tasks without retraining. Code and models will be made available.